6-2 Assignment: Addressing the Problem Through Sampling

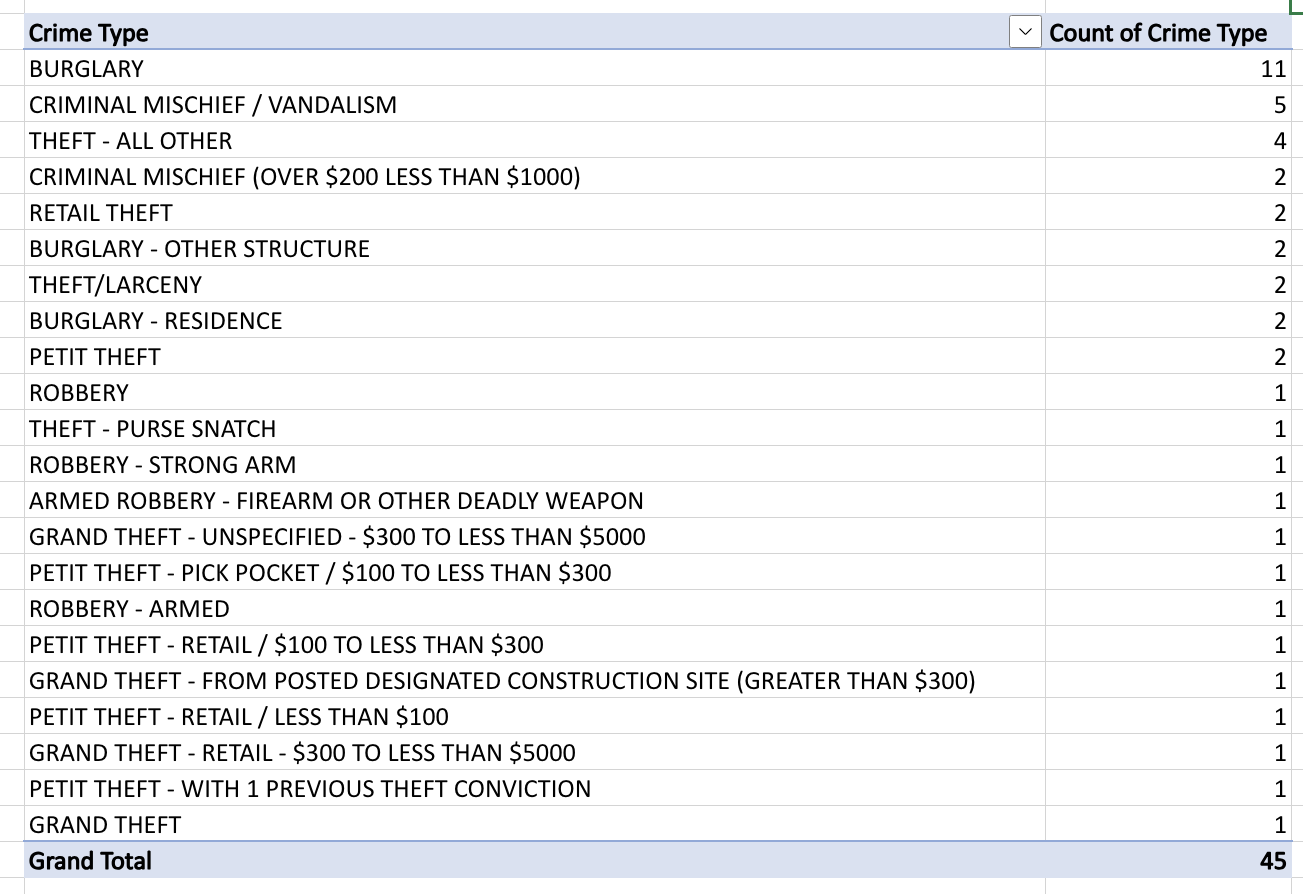
DAT-375

Ashley Littles

Southern New Hampshire University

I have been assigned with analyzing and sampling a large data set of crime types in Miami-Dade County. After analyzing the data in the larger data set, I found that the top three most frequent crimes are burglary, theft-all other, and criminal mischief/vandalism. When taking a sample from the larger data set, I found that the top three most frequent crimes are the exact same. Both analyses similarly have the same outcome even though I chose random samples for the second analyses. The only difference I would say is the size of the data sets. The sample set still has the same exact data as the larger data set but just a smaller amount. Even with a smaller amount of data in the sample set the three most frequent crimes are burglary, theft-all other, and criminal mischief/vandalism because overall they are the most common in the large data so when I select randomly every 10 rows, I will still have the same outcome. Taking a sample for analysis allows analysts to look at a larger data set without having to comb through the entire data set. In turn this saves data analysts time in the long run.





Reference

Bradburn, S. (n.d.) *How To Perform Random Sampling In Microsoft Excel* TopTipBio <https://toptipbio.com/random-sampling-excel/>